REMARKS/ARGUMENTS

Favorable reconsideration of this application, as presently amended and in light of the following discussion, is respectfully requested.

Claims 1-41 and 44-48 are pending, with Claims 2, 6, 20, 21, 25-28, 44, 47 and 48 amended and Claims 42-43 cancelled by the present amendment.

In the Official Action, Claims 2-8, 14, 17, 20, 22-23, 25-27, 38-39, 42-44 and 47-48 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Maxham et al. (U.S. Patent Publication No. 2004/0187075, hereinafter Maxham) in view of Koppich (U.S. Patent Publication No. US 2005/0141028). Claims 9 and 10 were rejected under 35 U.S.C. § 103(a) in view of Maxham, Koppich and Shaughnessy (U.S. Patent Publication No. 2004/0205664. Claim 11 was rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich and Howard (U.S. Patent No. 6,098,079). Claim 12 was rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich and Chi (U.S. Patent No. 5,978,917). Claim 13 was rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich, Howard and Kumashio (U.S. Patent Publication No. 2004/0193631). Claims 15 and 16 were rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich and Shaughnessy and Riss et al. (U.S. Patent Publication No. 2004/0103367). Claim 18 was rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich and Eagle (U.S. Patent No. 2003/0145209). Claims 19 and 21 were rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich and Howard. Claim 24 was rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich , Howard and Gladney (U.S. Patent No. 2003/0131241). Claims 28, 40, 41, 45 and 46 were rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich and Krachman (U.S. Patent No. 2004/0199555.). Claim 29-32 were rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich and Kenner et al. (U.S. Patent No. 6,421,726). Claims 33-34 were rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich ,Kenner and Black et al. (U.S. Patent Publication No. 2002/0059317. Claim 36 was rejected under 35

U.S.C. 103(a) in view of Maxham, Koppich, Kenner and McIver (CADINFO Article). Claim 37 was rejected under 35 U.S.C. 103(a) in view of Maxham, Koppich, Kenner, McIver and Windows (Windows Article)

Claim 2 is amended to recite Applicants' invention in varying language and to recite the features of now cancelled Claim 43. No new matter is added.

Briefly recapitulating, Claim 2 is directed to

> A method of profiling electronically-stored data, comprising:
>
> accessing data from one or more data archives, said one or more data archives located on one or more electronic archive devices;
>
> pre-filtering accessed data based on file content, content header information, and file meta-data, identified by the user to identify a pre-filtered file;
>
> transferring the pre-filtered file from said accessed data to a working electronic folder;
>
> tagging the working electronic folder with folder meta-data;
>
> processing the contents of said electronic folder, said step of processing including identifying whether a file within said electronic folder can be converted;
>
> selecting a file that can be converted in said processing step; and
>
> converting said selected file, said step of converting including
>
> extracting file meta-data from said selected file into a corresponding meta-data file,
>
> identifying a subset of text from said selected file;
>
> extracting only the subset of text from said selected file into a corresponding selected text file,
>
> creating an image of said selected file, and
>
> appending the corresponding meta-data file to said image of said selected file to create an appended image file.

Claim 48 is directed to a computer program product including instructions configured to control a computer so as to profile electronically-stored data, said instructions comprising

instructions for the method of Claim 2. Claim 47 is directed to as system including features similar to those of Claims 2 and 48.

Maxham describes a method for managing multiple documents to be uploaded to a document management computer system. The method includes determining a file type, creating a fingerprint for each document, deduplicating the documents in accordance with the fingerprint, extracting data from each document, associating extracted data with a corresponding document, and distributing the multiple documents and extracted data amongst a plurality of nodes of the document management computer system. Maxham also describes a method searching the multiple documents, a method for managing attributes of at least one of the multiple documents, and producing corresponding search results.[1]

In particular, Maxham describes that, before distributing input files 210 to computer system 10, file categorizer 216 creates a fingerprint of each file. Well known cryptographic algorithms, such as the MD5 check sum, may be used to create a fingerprint unique to each file. In accordance with the fingerprint, each document is deduplicated. Deduplicator 218 compares the fingerprint of each input file 210 with other fingerprints corresponding to other input files 210, and compares the fingerprint with the fingerprints of documents already stored in the computer system 10. If a match is found, the document is discarded, so as to prevent multiple documents from residing in the computer system 10.[2]

After the documents have been deduplicated, extractor 220 converts each native document 222 (corresponding to the input files 210 in original format) to at least a text file 224. The native document 222 may also be converted into a metadata file 226, XML file 228, and an HTML file 230.[3] Indexer 232 creates a file association table for each native document to maintain an association between converted documents 224-230, and any attachments (children files), to the native document.

[1] Maxham, paragraphs [0007]-[0011].
[2] Maxham, paragraph [0036].
[3] Maxham, paragraph [0037].

However, as acknowledged by the Official Action, Maxham does not disclose or suggest processing the contents of the electronic folder to include identifying whether a file within the electronic folder can be converted and selecting a file that can be converted in the processing step. To cure this deficiency, the Official Action applies Koppich.

Koppich describes a method for automatically performing preselected operations on electronic documents satisfying selected search criteria. The preselected operations of Koppich include: a) copying documents in the document data storage area to a selected destination; b) transferring documents to a selected destination; c) ***converting documents from a first selected file format to a second selected file format***; d) performing optical character recognition on a document to convert an image document to text; e) assigning selected values to metadata fields; and f) identifying criteria which must be associated with a document in order fro the preselected operation to be performed on the document.[4]

Koppich further describes that the document management system includes means for determining whether it is possible to execute the selected operations or script. If the determination is positive, the selected operations are stored in a suitable memory device. If the determination is negative, the user is prompted to edit or revise the script. The revised script is then reviewed to determine if the selected operations are able to be executed.[5] In a "convert to" operation, documents received in the data storage area are converted from a first selected file format to a second selected file format. To select the appropriate formats, the user provides information about the formats. Koppich describes the documents are converted from PS to PDF, from PDF to PS, from PS to TXT, from PDF to TXT, from PS to TIFF, from PS to JPG, and from PS to BMP. Koppich notes that other types of document conversion may be possible.[6]

[4] Koppich, paragraphs [0008]-[0011].
[5] Koppich, paragraph [0044].
[6] Koppich, paragraph [0050].

However, Maxham and Koppich each do not disclose or suggest selecting a subset of text from an entirety of text of the selected file via a user interface. That is, Maxham and Koppich each do not disclose or suggest "identifying a subset of text from said selected file" or "extracting only the subset of text from said selected file into a corresponding selected text file" as recited in Claims 2, 47 and 48. By being able to extract only a user-selected subset of text, an end-user is able to more precisely extract text for downstream processing, and to exclude text that is either irrelevant or private.

Maxham and Koppich also fail to disclose or suggest prefiltering based on file content, content header information and file metadata, previously recited in now cancelled Claim 43. In rejecting Claim 43, the Official Action notes that paragraph [0035] of Maxham describes a file type discriminator 212 that determines file types based on a file extension of each input file 210. Only files that match a certain file type are extracted by archive extractor 214. However, the file type extractor of Maxham does not include a prefilter that filters based on file content, content header information and file metadata. With Applicants' claimed prefiltering based on file content, content header information and file metadata, more efficient screening of documents is possible.

Applicants have considered the remaining documents applied in the outstanding rejections, and submit that none of the remaining applied references cure the deficiencies of Maxham and Koppich.

As none of the cited prior art, individually or in combination, discloses or suggests all the elements of independent Claims 2, 47 and 48, Applicants submit the inventions defined by Claims 2, 47 and 48, and all claims depending therefrom, are not rendered obvious by the asserted references for at least the reasons stated above.[7]

[7] MPEP § 2142 "...the prior art reference (or references when combined) must teach or suggest **all** the claim limitations.

Regarding dependent Claim 20, Maxham and the remaining references fail to disclose or suggest the steps of "creating a ***single*** searchable ***master text file containing the contents of all selected files***; time stamping or digitally authenticating the searchable master text file; and appending selected meta-data about the files included in the master text file." The cited portion of Maxham used to reject Claim 20 describes that "***each*** of a plurality of documents having at least one of either meta-data, text or attachments identified for retrieval that are indexed for web-based retrieval from the ***cluster database***, said identification of the plurality of documents forming a cluster data base that is web-searchable by use of a predetermined descriptive term." As noted above, Maxham also discloses that "after the documents to be uploaded have been de-duplicated, extractor 220 converts each native document 222 (corresponding to the input files 210 in original format) to at least a text file 224." However, plural text attachments, or the creation of multiple text files 224 (i.e., one for each native document 222) is not equivalent to creating a ***single*** searchable ***master text file containing the contents of all selected files***. By having a single master text file, searching across an entire library of documents is more easily facilitated. Thus, for independent reasons, Applicants submit that Claim 20, and all claims depending therefrom, patentably define over the applied references.

Accordingly, in view of the present amendment and in light of the previous discussion, Applicants respectfully submit that the present application is in condition for allowance and respectfully request an early and favorable action to that effect.

Respectfully submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

Bradley D. Lytle
Attorney of Record
Registration No. 40,073

Michael E. Monaco
Registration No. 52,041

Customer Number
22850

Tel: (703) 413-3000
Fax: (703) 413 -2220
(OSMMN 08/07)
MM/rac
I:\ATTY\MM\250727US-AM.DOC